



# Processing 3D form and 3D motion: Respective contributions of attention-based and stimulus-driven activity

A.-L Paradis, J Droulez, V Cornilleau-Pérès, J.-B Poline

## ► To cite this version:

A.-L Paradis, J Droulez, V Cornilleau-Pérès, J.-B Poline. Processing 3D form and 3D motion: Respective contributions of attention-based and stimulus-driven activity. *NeuroImage*, 2008, 10.1016/j.neuroimage.2008.08.027 . hal-01230996

**HAL Id: hal-01230996**

**<https://hal.science/hal-01230996>**

Submitted on 19 Nov 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **Processing 3D form and 3D motion: respective contributions of attention-based and stimulus-driven activity.**

A.-L. Paradis<sup>1,4</sup>, J. Droulez<sup>2,4</sup>, V. Cornilleau-Pérès<sup>2</sup>, J.-B. Poline<sup>3,4</sup>

<sup>1</sup> CNRS, UPR640, Laboratoire de Neurosciences Cognitives et Imagerie Cérébrale, 75013 Paris, France;

<sup>2</sup> CNRS – Collège de France, UMR7152, Laboratoire de Physiologie de la perception et de l'action, 75005 Paris, France;

<sup>3</sup> Neurospin, I2BM, CEA, Bâtiment 145, 91191 Gif-sur-Yvette cedex, France ;

<sup>4</sup> IFR49 Institut d'Imagerie Neurofonctionnelle

Running title: stimulus and attention in 3D-SFM perception

Contact: Anne-Lise PARADIS

LENA, CNRS – UPR640

47, bld de l'hôpital, 75651 Paris cedex 13

France

Email: anne-lise.paradis@chups.jussieu.fr

Fax: 33-1 45 86 25 37

**Keywords:** fMRI; structure from motion; selective attention; stimulus-driven processing; 3D shape; 3D motion; dorsal pathway; ventral pathway.

## **ABSTRACT**

This study aims at segregating the neural substrate for the 3D-form and 3D-motion attributes in structure-from-motion perception, and at disentangling the stimulus-driven and endogenous-attention-driven processing of these attributes.

Attention and stimulus were manipulated independently: participants had to detect the transitions of one attribute –form, 3D motion or colour– while the visual stimulus underwent successive transitions of all attributes. We compared the BOLD activity related to form and 3D motion in three conditions: stimulus-driven processing (unattended transitions), endogenous attentional selection (task) or both stimulus-driven processing and attentional selection (attended transitions).

In all conditions, the form versus 3D-motion contrasts revealed a clear dorsal/ventral segregation. However, while the form-related activity is consistent with previously described shape-selective areas, the activity related to 3D motion does not encompass the usual “visual motion” areas, but rather corresponds to a high-level motion system, including IPL and STS areas.

Second, we found a dissociation between the neural processing of unattended attributes and that involved in endogenous attentional selection. Areas selective for 3D-motion and form showed either increased activity at transitions of these respective attributes or decreased activity when subjects’ attention was directed to a competing attribute. We propose that both facilitatory and suppressive mechanisms of attribute selection are involved depending on the conditions driving this selection. Therefore, attentional selection is not limited to an increased activity in areas processing stimulus properties, and may unveil different functional localization from stimulus modulation.

## **ABBREVIATIONS**

3D = three-dimensional; SFM = structure from motion; BOLD = blood oxygenation level dependent; (f)MRI = (functional) magnetic resonance imaging; ROI = region of interest.

## INTRODUCTION

Visual motion is a rich source of information about the environment: from motion cues only, we are able to perceive our self-motion (direction of heading), other's actions (biological motion) and, of primary interest in this study, the 3D structure and 3D motion of the surrounding objects.

Structure-from-motion (SFM) perception has been largely demonstrated and tested using dynamic random dot stimuli, for which the motion parallax (i.e. the relative motion between dots) is the only depth cue (Rogers and Graham, 1979; Braunstein and Andersen, 1984; Cornilleau-Pérès and Droulez, 1994). The neural substrate of SFM perception has been explored in various imaging studies (Orban et al., 1999; Paradis et al., 2000; see also the review by Greenlee et al., 2000; Kriegeskorte et al., 2003; Murray et al., 2003; Peuskens et al., 2004). Overall, optical flows generating SFM perception activate a large set of visual areas, not specific to the extraction of the structure information from motion: this SFM network includes the visual motion areas (including V2, V5+ and regions of the intraparietal sulcus); ventral areas involved in shape perception (lateral occipital and fusiform cortices; collateral sulcus) and areas presumably involved in the control of attention (in the intraparietal and precentral sulci). Our goal is to better understand the respective role of these visual and attentional areas in the processing of two different “end-products” of SFM perception: the 3D form and its 3D motion.

### *3D structure and 3D motion from 2D motion*

The perception of 3D motion is a correlate of SFM perception. This is well illustrated by the simultaneous alternation of motion direction together with 3D shape in the bistable perception of a rotating Necker's cube. This was also demonstrated mathematically by Longuet-Higgins and Prazdny (1980), who established that the 3D movements and 3D structure are recovered altogether through the same process. Extracting the 3D movements of a visual stimulus from the retinal 2D motion indeed requires non trivial processing: translations on the retina, for instance, may correspond to the rotation of a 3D stimulus around a fronto-parallel axis. Yet, little interest has been devoted to the perception of “3D motion from 2D motion” compared to the perception of structure from motion. The first aim of the present study is to disentangle the respective contributions of 3D form and 3D motion perception to the cerebral activity induced by an optical flow.

### ***One input, two visual pathways***

Although intimately associated in the optical flow, the form and 3D motion of the underlying objects are well segregated at the perceptive and physiological levels. Structure and motion direction are easily identified as two distinctive attributes of a perceived object. Form conveys information about the identity of the object while movements usually do not, even if motion has also been explored as an intrinsic property of objects (see Newell et al., 2004 and the concept of spatio-temporal signature by Stone, 1998). Importantly, 3D-form and 3D-motion attributes can vary in an independent way.

From a physiological viewpoint, form and motion are processed along two distinct visual pathways. Form processing is carried out by the ventral pathway devoted to object identification, whereas motion processing develops along the dorsal pathway devoted to visuo-spatial interactions (Ungerleider and Mishkin, 1982; Goodale, 1998). Accordingly, SFM perception should activate both the ventral and dorsal pathways.

Previous studies exploring the neural bases of SFM perception showed that both pathways were indeed activated differentially when comparing a 3D-SFM stimulus to a non-coherent 2D-motion display (Orban et al., 1999; Paradis et al., 2000; Kriegeskorte et al., 2003; Murray et al., 2003). These results indicate that visual processing within the ventral path is not limited to static cues, and that the dorsal path does not exclusively process motion information. However, these studies did not fully elucidate the respective roles of the ventral and dorsal pathways in SFM perception. Are the ventral and dorsal activities related to early-processing stages (e.g. retinal-speed analysis, extraction of depth information...); are they related to the processing of various perceptual attributes (form and 3D motion of the visual object); or do they reveal tasks implicitly performed on the object (e.g. identification, simulated manipulation, orientation judgment, etc.)?

### ***Respective contributions of stimulus-driven and attention-related processes?***

To better control the possible influence of an implicit task and disentangle the respective contribution of form and motion attributes on SFM-related activity, several authors introduced a task to focus subjects' attention on different attributes of the 3D object<sup>1</sup>. Activity was found

---

<sup>1</sup> Following Corbetta and other's results, the working hypothesis is that selective attention to a visual attribute enhances the activity in areas processing this attribute (Corbetta *et al.*, 1991; Huk and Heeger, 2000). Hence, comparing conditions where subjects attended to the 3D form versus conditions where their attention was focused on the direction of motion was expected to highlight the areas specialized in 3D-form processing with respect to those specialized in 3D-motion processing.

predominantly in the dorsal pathway when observers attended to the direction of motion and predominantly in the ventral pathway when observers attended to the form or texture (Paradis et al., 2001; Peuskens et al., 2004). While informative, those studies tested the effect of feature-directed attention only, averaging BOLD activity over different visual conditions. Yet, different mechanisms may take place depending on whether the stimulus itself remains identical or changes over time.

In the present work, we clarify the contribution of attention-related and stimulus-driven inputs to the processing of 3D motion and structure in SFM perception. To disentangle the stimulus-driven processing from attentional selection, we independently manipulated the physical attributes of the stimulus and the participants' attention. The stimulus underwent changes of form, direction of 3D motion and colour distribution. Meanwhile, observers' attention was focused by a detection task: a visual instruction prompted them to report the transitions of either form, 3D motion or colour distribution until the next instruction.

To characterize stimulus-driven activity, we tested the effect of the form and 3D-motion transitions while participants were attending to the colour changes. In the following, these transitions are called "unattended transitions". To characterize attention-related activity, we tested, at the colour transitions, the influence of attending to form or to 3D motion. Last, to evaluate the contribution of selective attention to the visual processing of form and 3D motion, we analysed the activity elicited by the form transitions when subjects were attending to form, or by the 3D-motion transitions when subjects attended to 3D motion. These are called "attended transitions" in the following.

## **METHODS**

### ***Participants***

Eleven healthy volunteers (5 men and 6 women) aged 21-28 years took part in the study, approved by an Institutional Ethic Committee (CCPPRB). Volunteers gave their written informed consent and received a small financial compensation for their participation.

All participants had normal vision; all but one were right handed; one had a left ocular dominance. One subject was excluded from the analysis because of excessive head movements (above 3 mm displacement in translation).

### ***Visual stimuli: SFM with transitions of form, 3D motion, and colour distribution***

The visual stimuli, presented over a black background, comprised a central fixation cross and a distribution of 200 coloured dots (red and green antialiased dots, 6 pixels width, 0.27° visual angle; perceptual equiluminance between red and green was achieved for each participant using an equalisation procedure based on heterochromatic flicker photometry). During the stimulation, the dots continuously moved as if they belonged to

a 3D surface oscillating in depth around a fronto-parallel axis tangent to the surface (sinusoidal oscillation: 10° maximal amplitude, 2 sec period; see Paradis et al., 2000). This stimulus was viewed through a 16° diameter virtual window; moving dots could appear or disappear behind the invisible edges of this mask, but the edges of the surface were never visible.

Every 2 seconds, when the oscillating surface passed through the central (and initial) position, either the form, the orientation of its oscillation axis or the distribution of dot colours could change: the 3D form alternated between a paraboloid and a horse-saddle; the oscillation axis could tilt in the screen plane by an angle of 45°, 60° or 90°; part of the dot distribution (85 to 95 %) could reverse colour from red to green or vice versa. The order of the transitions (form, 3D-motion direction and colour change) was randomized.

The 3D parameters of the stimulus –its surface curvature and oscillation amplitude– were chosen so that all motion and form transitions yielded a similar amount of visual acceleration. Because of the surface movement, this visual acceleration was always minimal at the centre of the screen. In order to minimize the visual change at the centre of the screen for the colour transitions as well, no dot under 1° eccentricity changed colour. Also, the percentage of dots changing colour was adjusted so that the participants could achieve similar performances in the colour-related task (see below) compared to the form- and motion-related tasks (preliminary psychophysical experiments over 9 subjects, not shown).

#### *Task: detecting the transitions of one attribute of the stimulus*

Experimental runs consisted of 9 stimulation blocks (52 sec each) separated by instruction screens (one word presented for 2 sec). Participants were instructed to fixate the central cross and press a button when detecting transitions of the attribute designated by the instruction screen ("form", "motion" or "colour"). Instruction screens were used as a low-level baseline in some analyses. Participants underwent three training runs before entering the scanner, and two runs while being scanned.

#### *Experimental set-up*

In the scanner, stimuli were back-projected on a translucent screen using an Eiki 5000 projector driven by a personal computer; participants could see the screen located at the head-end of the magnet through mirror glasses. Participants' responses were collected through a non-magnetic push-button held in the dominant hand. The computer was connected to the MRI scanner so that stimulation and subjects' responses could be precisely measured with respect to the acquisition time course.

After completion of the scanning sessions, participants were asked to report on their different visual percepts while they were performing the experiment.

#### *MRI acquisition*

Participants were scanned, using a 3T whole body MRI system (Bruker, Germany). Functional data were obtained with a T2\*-weighted gradient echo EPI sequence (flip angle 90°, TE = 40ms, TR = 2 s) sensitive to BOLD contrast. Each volume comprised 18 contiguous slices (in-plane resolution 3.75 x 3.75 mm<sup>2</sup>, 6 mm thickness) generally covering the cerebral cortex, but excluding the cerebellum and the most inferior part of the temporal poles. In the scanning session, participants underwent four runs of functional acquisition; only two are relevant to the present work. A high resolution (1.5 x 2 x 1 mm<sup>3</sup>) T1-weighted IR gradient echo sequence

(Inversion Time  $TI = 700$  ms,  $FOV = 192 \times 256 \times 256$ ) was also performed to acquire accurate structural information.

### ***Data analysis***

Functional data were first corrected for their geometric distortions, using a home-made unwarping procedure based on the characterisation of field inhomogeneities described by Jezzard and Balaban (1995).

Except for one subject, excluded from the analysis, the evaluated amount of head movement within runs was less than 3 mm in translation and  $2^\circ$  in rotation. To avoid a possibly prejudicial correction (Freire and Mangin, 2001) realignment was performed between but not within runs. We assumed that subjects' movements were negligible between the run used as the target of the realignment and the acquisition of the structural image.

Further pre-processing was performed using the standard SPM procedures (<http://www.fil.ion.ucl.ac.uk/spm>). Anatomical and functional images were normalised into the MNI stereotactic system of coordinates by fitting the anatomical images to a local template that matches better the contrast specificity of our images, and applying the subsequent linear and non linear transformations to the realigned images. Functional individual data were smoothed (8-mm isotropic Gaussian kernel full width at half maximum) to facilitate the coregistration at the group analysis level. The voxel size of the normalised functional volumes was set to  $3 \times 3 \times 3$  mm<sup>3</sup>. The first three scans acquired during the transition to the steady state of the magnetic resonance signal were discarded.

Statistical analyses were carried out using SPM2 (<http://www.fil.ion.ucl.ac.uk/spm>). Individual results were entered in a second level random effect analysis to obtain group results. Statistical inferences are based on  $t$ -statistics over the estimated parameters of the model, converted into  $z$ -scores. Areas of interest were selected using a double threshold of  $p < 0.01$  uncorrected over the voxel and  $p < 0.05$  uncorrected over the cluster. Only activation maxima located inside these areas are reported (see tables).

### ***Model of the individual BOLD response and contrasts of interest***

Nine regressors were built-up using the “HRF” basis function triggered by the stimulus transitions<sup>2</sup>: each regressor modelled the event-related BOLD responses to one type of transition (i.e. change of 3D motion, form or colour) under one type of task (i.e. detection of 3D motion, form or colour transitions).

Six contrasts were calculated in order to compare the neural substrate recruited for 3D motion and form processing (“form minus 3D motion” and “3D motion minus form”), when the activity was (1) stimulus driven; (2) attention based or (3) induced by a combination of visual and attentional inputs: the colour-related task was used to compare the responses to the form and motion transitions in the context of an incidental task (i.e. “unattended transitions”, see contrasts of stimulus-driven activity (1) in Figure 1); the colour transitions allowed comparing the conditions of form- and motion-directed attention, independently of the stimulus changes (see contrasts of attention-related activity (2) in Figure 1); we also compared the form and 3D motion transitions

---

<sup>2</sup> We also tested a model of the detection events based on the response of the subjects, slightly anticipated (500ms) to account for the delay between the neural response and the button click. Response-based regressors were very similar to the stimulus-based ones (calculated for 10 subjects \* 3 conditions \* 2 runs, correlation coefficients range from 0.48 to 0.97, with 91% of the values over 0.8), and the results did not differ qualitatively from the stimulation-based model.



occurring while the participants attended to these attributes ( “attended transitions”, see contrasts of stimulus-driven and attention-related activity (3) in Figure 1).

We used the SPM toolbox MarsBaR (Brett *et al.*, 2002; <http://marsbar.sourceforge.net/>) to test the average responses of the regions delineated by the group results to each type of transition.

Figure 1 about here

## RESULTS

### *Behavioural results*

For each task, we count a correct response when the subject presses the button between 100ms and 2s after the transition to be detected. The detection rate is the ratio between correct responses and the total number of transitions to detect. Group results were calculated over the 10 participants included in the fMRI analysis.

According to the preliminary tests, the detection rate did not vary significantly between tasks ( $91 \pm 10 \%$ ,  $93 \pm 9 \%$  and  $96 \pm 6 \%$ , for the colour, form and motion tasks respectively; all  $p > 0.1$  for paired t-tests between individual detection rates; see Figure 2). Reaction times for the form and motion tasks were not significantly different either (Figure 2; mean =  $1071 \pm 153$  ms and  $1044 \pm 102$  ms respectively;  $p = 0.5$  for the bilateral paired t-test). Overall, the motion and form tasks yielded similar behavioural performances.

The detection of colour changes (mean latency =  $849 \pm 159$  ms) was significantly faster than the detection of form and 3D-motion changes ( $p < 0.002$  for both paired t-test). The shorter reaction time was not correlated with a subjective feeling of easiness, since 6 out of the 10 participants ranked the colour task as the most difficult; in comparison, the motion and form tasks were considered as the most difficult by 1 and 3 participants, respectively. However, we never directly compare the responses to the colour task with the responses to the motion or form tasks, so that the difference in reaction times cannot be a possible confound.

Figure 2 about here

After the scanning sessions, participants were further asked whether they had noticed other transitions than those they were supposed to detect during each task. For the colour task, 6 participants reported not seeing any unattended transition of either form or 3D motion, 1 noticed form transitions and 3 noticed both form and 3D motion transitions. During the motion task, 6 participants could see form transitions, 2 noticed colour transitions and 3 did not notice anything but the attended transitions of motion. During the form task, 6 participants

could see motion transitions, 1 noticed colour transitions and 3 did not see anything but the attended transitions of form.

Participants tend to notice form (resp. 3D motion) transitions less often during the colour task than during the 3D motion (resp. form) task, as if the colour task drew the participants' attention away from both form and motion information. This supports the working hypothesis of the colour task being incidental to the form and motion transitions. Moreover, the colour task does not significantly favour the perception of one type of transitions (form or 3D motion) compared to the other. Thus, we do not expect unspecific activity related to distractor detection when comparing the unattended transitions of form and 3D motion. In the same way, participants did not perceived more colour transitions during the form task than during the motion task (or vice versa), so we do not expect activity due to an unbalanced detection of the colour transitions when comparing the form and motion tasks at these transitions.

### ***Imaging results: 3D form vs. 3D motion***

In the following, we compare the neural substrates recruited to process 3D motion and 3D form when the activity is (1) stimulus driven; (2) attention related or (3) induced by a combination of visual and attentional inputs.

#### *(1) Stimulus-driven activity*

Figure 3-1 shows the areas responding differentially to the unattended form transitions and unattended 3D-motion transitions (same incidental task, but different transitions).

The main result is a clear ventral/dorsal segregation, with occipito-temporal activity selective for 3D shape ("form minus 3D motion" contrast) and parieto-frontal activity selective for motion direction ("3D motion minus form" contrast).

Foci more activated by the form than by the 3D motion transitions were found bilaterally along the superior occipital sulcus (corresponding to V3/V3a), and in the inferior temporal gyrus (see Table 1-A). These areas are part of the network activated by SFM stimuli in passive viewing conditions. It is noteworthy that the activity in the inferior temporal gyrus is close to the V5+ complex (as defined by an independent localizer experiment with the same subjects), but the two regions do not overlap. From its coordinates and its anatomical localization, the focus of activity could correspond to the LOC (Lateral occipital complex: Grill-Spector et al., 1998).

Reciprocally, areas selective for 3D-motion transitions were found in the superior and middle frontal gyri and in both inferior parietal lobules (more precisely in the supra-marginal gyri,

see Table 1-B). The middle frontal focus could correspond to the “lateral” frontal eye fields as described by Grosbras et al. (2001). In contrast to the shape selective foci, the areas we find selective for 3D motion do not correspond to the classical “visual motion” areas and are not usually reported as part of the SFM network.

### *(2) Attention-related activity*

To highlight the activity related to the attentional selection of the 3D attributes, we compared the influence of attending to form or to 3D motion at the colour transitions (same transition, different tasks).

Only the 3D motion minus form contrast showed significant activation. Areas selectively activated by the attention to 3D motion were found in the left middle temporal gyrus and bilaterally in the inferior parietal lobules (see Figure 3-2). Even if the parietal foci are distinct from the stimulus-driven activity described above (coordinates of the local maximum in the left inferior parietal lobule are respectively -45 -78 33 and -51 -48 48; see Table 1 and Table 2), we thus find that the inferior parietal lobule is involved both in the stimulus-driven processing and the attentional selection of the 3D motion attribute. This result agrees with the dorsal localisation of 3D-motion processing.

### *(3) How do stimulus-driven and attention-related activities combine?*

The comparison of the attended form transitions and attended 3D-motion transitions revealed a segregation similar to that observed with the unattended transitions: a ventral occipito-temporal network for the “form minus 3D motion” contrast and a dorsal parieto-frontal distribution of activity for the “3D motion minus form” contrast (see Figure 3-3). However, the activity evoked by the attended transitions and that evoked by the unattended transitions did not overlap.

The attended transitions revealed form-selective areas in the temporo-occipital cortex, and along the collateral sulcus (see Table 3-A). Compared to these regions, the stimulus-driven foci were located posteriorly, suggesting that the unattended form transitions involved earlier visual areas.

The 3D-motion minus form contrast of attended transitions yielded mesial activity that was not highlighted by the contrast of unattended transitions. This new 3D-motion-selective activity lies around the supplementary motor area (paracentral lobule) and in the precuneus, both along the parieto-occipital sulcus and the posterior cingulate (see Table 3-B).

Besides, attended transitions also revealed 3D-motion-selective activity close to regions highlighted by the unattended transitions. This activity lies in the bilateral inferior parietal

lobules and in the left superior frontal sulcus. On the right hemisphere, the superior frontal focus was close to a significant cluster threshold ( $p=0.053$ ). The coordinates of these frontal foci differed from those found with the unattended transitions (respective local maxima being 19 mm apart on the left and 18 mm apart on the right). They could correspond to the dorsomedial aspect of the frontal eye field, whereas the regions for the unattended transitions corresponded better to the lateral FEF (see Table 1 and Table 3).

The activity of the inferior parietal lobule related to the attended transitions was found in the angular gyrus. This location does not correspond to the foci evoked by the unattended transitions, but it is largely superimposed with the foci evoked by the attention-related contrast. This suggests that the angular gyrus is involved in the attentional selection of 3D-motion, whether a stimulus-driven processing is engaged or not.

Figure 3 about here

*To summarize, all three contrasts show a clear dorsal/ventral segregation of the neural substrates for the processing of the 3D motion and form attributes. The ventral form-selective areas are part of the network activated by the passive viewing of SFM stimuli, but the dorsal areas, selective for 3D motion, fall outside the SFM network. More intriguing, the areas activated by attended transitions are spatially distinct from those activated by unattended transitions.*

***Are the areas recruited by the attended transitions functionally different from those activated by the unattended transitions?***

The question is specifically relevant for the regions that are close to each other in the two contrasts (unattended and attended). Do we find different foci because of an intrinsic variability of the localization in the group results, or do these foci correspond to functionally different areas? To answer this question, we further analysed the BOLD activity within these areas. We tested whether this activity was increased or decreased with respect to a low level baseline<sup>3</sup> in response to the transitions (see figure 4). We also tested, separately for each attribute, the difference between the conditions driving the attribute processing (unattended transitions, task and attended transitions). Unless specified, those contrasts are orthogonal to those used to define the region of interest.

---

<sup>3</sup> The baseline consisted of a blank screen with the instruction word (see Methods).

### *Shape-selective areas around the collateral sulcus*

We tested the bilateral region corresponding to the LOC (upper boxes in Figure 4-B). By definition, this region is more activated by the unattended form transitions than by the unattended 3D-motion transitions. Analyses show that this difference is due to a significant increase of activity at the unattended form transitions. Independent of the unattended transitions, the activity at the attended form transitions is also increased, but no activity is found during the form task (left group of 3 (red) bars). The difference of activity between the attended form transitions and the form task is significant ( $p=0.03$  on the left and  $p=0.02$  on the right). This means that this region is sensitive to the form transitions not only in the context of the colour task (unattended transitions) but also in the context of the form task. Altogether, the LOC seems sensitive to form transitions independently of the task context.

The most anterior regions reveal a very different behaviour. By definition, this region is more activated by the attended form transitions than by the attended 3D-motion transitions. This significant difference however is not due to an increased activity at the attended form transitions, as it could be expected, but corresponds to a decreased activity at the attended 3D-motion transitions (see orange boxes in Figure 4-B). In line with this, these anterior regions do not reveal any response to the unattended form transitions ( $p>0.3$ ), but show a significantly negative response during the motion task ( $p<0.001$ ). Hence, contrary to the LOC, the shape-selective anterior regions are not sensitive to the form transitions. Instead, they reveal a strongly negative signal during the motion task, independent of stimulus transitions.

### *3D Motion selective areas in the superior frontal gyrus*

The dorsomedial FEF (see orange boxes in Figure 4-A) was delineated by the significant difference of activity between the attended 3D-motion transitions and the attended form transitions. Analyses show that this difference reflects an increased activity at the attended 3D-motion transitions with respect to the low level baseline ( $p=0.004$  on the left and  $p=0.001$  on the right). Independent of the attended transitions, the unattended 3D-motion transitions also induce increased activity (not significant on the left but  $p=0.02$  on the right). Eventually, the activity is decreased similarly during the form task and the 3D-motion task at the colour transitions. Overall, the dorsomedial FEF appears sensitive to the 3D-motion transitions whatever the task context is, but shows a stronger BOLD signal during attended transitions.

The lateral FEF (blue boxes in Figure 4-A) was delineated by the significant difference of activity between the unattended 3D-motion transitions and the unattended form transitions. This difference corresponds to an increased activity at the unattended 3D-motion transitions (although not significant on the right). Paradoxically, the analyses do not show any significant

response compared to the baseline when the 3-D motion transitions are attended. Besides, the lateral FEF shows significantly decreased activity at the attended form transitions (left and right  $p=0.003$ ) and during the form task ( $p\leq 0.001$ ). Hence, the region is both negatively modulated by the form task, independently of the stimulus transitions and activated by the unattended 3D-motion transitions only.

Figure 4 about here

#### *Inferior parietal lobule*

In the inferior parietal lobule, we compared the activity of the supramarginal gyrus found in the contrast of unattended transitions with that of the angular gyrus found in the contrast of attended transitions. Both regions showed increased activity with respect to the baseline at the unattended 3D-motion transitions, but no significant activity at the attended 3D-motion transitions. The two of them also showed decreased activity during the form task, either at the colour transitions or at the attended form transitions, but no significant activity at the unattended form transitions. We conclude that the whole inferior parietal lobule globally follows the same pattern of activity, similar to that of the lateral FEF.

#### *Complementary results*

A similar analysis was conducted in the regions of interest having no obvious counterpart in the other contrasts (such as the superior occipital focus, see table 1). The results are reported as supplementary material.

*The results of this section confirm that the areas delineated by the contrast of unattended and attended transitions, both on the ventral and superior frontal cortices, correspond to functionally distinct areas. In the inferior parietal lobule, however, the supramarginal and angular foci cannot be segregated on functional criteria.*

## **DISCUSSION**

In the following, we first examine how the visual motion and shape selective areas of the classical SFM network behave in our paradigm and discuss their respective contribution to the processing of 3D-motion and form. We then examine the specific neural substrate processing 3D motion. Lastly, we discuss the differences observed between unattended and attended transitions. We shall specially consider how these observations unveil the substrate of attentional selection mechanisms for the form and 3D-motion attributes studied here.

### ***Modulation of the visual motion and shape selective areas by the processing of 3D form and motion***

Previous studies used different viewing situations to delineate the cortical areas involved in SFM perception (passive viewing: Orban et al., 1999; Paradis et al., 2000 and 2001; object recognition: Kriegeskorte et al., 2003). From these, a common SFM network can be described, which includes ventral areas selective for the shape and the “visual motion areas” (V5+, V3/V3A and intraparietal motion sensitive areas). In the following, we discuss the activity of these specialized areas relative to the form and motion attributes in our paradigm.

#### *Lateral occipital and temporo-occipital cortex*

Unattended transitions highlighted shape selective areas in the lateral occipital cortex, which are likely to correspond to the LOC (Grill-Spector et al., 1998; Kourtzi et al., 2003). Indeed, the sensitivity of the region to the form transitions, either attended or unattended, is consistent with the selectivity of the LOC to 3D shapes (Kourtzi et al., 2003). Sensitivity to transitions is also a logical counterpart to the strong adaptation to repeated shapes that has been described in this region (Grill-spector et al., 1999).

Attended transitions revealed more anterior foci, lying on the ventral aspect of the temporal lobes, at the borders of the SFM network. We checked that, in contrast with the LOC, these foci were not sensitive to the form transitions but showed decreased activity during the 3D-motion task. Their location corresponds to areas specialized in shape categories (cf. Martin et al., 1996; Ishai et al., 1999). We thus observe that more specialized areas are recruited by the attended transitions and modulated by the attention to a competing object attribute, while early visual areas are rather activated by physical stimulus changes. This suggests that the mechanisms of attribute selection involved at different stages of the visual hierarchy are different.

#### *Superior occipital cortex*

The superior occipital region (including the junction of the intraparietal and intraoccipital sulci, V3A, and part of the lateral occipital gyrus) was selective for the unattended form transitions. We checked that this region was indeed activated by the unattended transitions of both 3D motion and form (see supplementary material, Figure 5-A), but reached maximal activity for the form transitions. These results are fully consistent with previous studies, as V3A is known to be sensitive not only to visual motion (Tootell et al., 1997) but also to shape (Denys et al., 2004). Moreover, the superior occipital region already revealed a particular sensitivity to the 3D content of a visual stimulus and to its curvature in passive viewing,

suggesting a role in the analysis of the optic flow and the extraction of the 3D structure (Paradis et al., 2000).

More precisely, this region may extract the orientation of the object principal axis. Recently, Valyear et al. (2006), using static objects, highlighted an area at the occipito-parietal junction (OPJ) that closely corresponds to our right superior occipital focus (see coordinates in table 1), and is sensitive to changes of orientation of the object principal axis. This may seem at odds with our results because the region we describe is more sensitive to the transitions of form than to the transitions of motion direction. Our 3D-motion transitions however do not modify the orientation of the object principal axis, whereas our form transitions correspond to changes of the curvature axes, which are also the principal axes of the object. Overall, the superior occipital region (area OPJ) seems to be involved in the automatic extraction of coarse information about the global structure of 3D objects.

#### *V5+ complex*

Although the activity in the V5+ complex was enhanced at the form and motion transitions in all tasks (see supplementary material, Figure 5-A and C), V5+ was not significantly modulated by the attribute attended to by the participants. This is congruent with studies showing that V5+ activity is hardly modulated by the task, particularly when speed or motion direction is concerned (Cornette et al., 1998; Sunaert et al., 2000). Why, however, did other studies show significant modulation of V5+ activity with the selective attention to motion in monkeys and humans? (see Treue and Maunsell, 1996; Büchel et al., 1998; Chawla et al., 1999). The difficulty level does not seem to account for this as difficult tasks may induce either strong or weak modulation in V5+. Rather, the modulation of V5+ activity consistently depends on whether attention is directed toward the visual motion input or diverted from it. In the present work, the activity in V5+ for the form and motion tasks confirms that the visual motion input is mandatory for both tasks and suggests that V5+ provides a common source of information for processing form and 3D-motion.

During the colour task (selective attention to colour), part of the right V5+ appears more activated by the form transitions than the 3D-motion transitions. This observation is consistent with V5 showing selectivity to depth gradients in the monkey (tilted planes in Xiao et al., 1997). However, the lack of modulation by attention suggests that, despite this selectivity, V5+ is not directly involved in detecting 3D-form transitions. More surprising, the absence of activation specific to 3D-motion (either through physical transitions or selective attention) suggests that V5+ is not directly involved in detecting changes of 3D motion either. These results, however, are not so paradoxical if we consider that V5+ implement a stage of



2D visual motion processing, providing a common source of information for processing 3D motion and form.

#### *Posterior-parietal Cortex*

As V5+, the posterior parietal cortex is activated by attention to visual motion (Büchel et al., 1998), is activated by the unattended transitions of form and 3D motion in the present study (see supplementary material, Figure 5-A and 5-C) and is not selective for 3D motion or form. A major difference with V5+, however, is that the attended form and motion transitions do not activate this region (see supplementary material, Figure 5-A and 5-C). This indicates that the posterior parietal cortex does not process all visual motion changes as does V5+. Instead, this result is consistent with a role of the posterior parietal cortex in engaging visual attention (Corbetta and Shulman, 2002): the activity of the region at the unattended transitions could correspond to an exogenous attraction of the attention toward the visual motion input. This region however is not a candidate substrate for the processing of 3D motion.

#### ***3D-motion attribute: visual motion or not visual motion?***

The SFM network largely overlaps the cortical areas sensitive to visual motion (Dupont et al., 1994) and to the attention to motion (Büchel et al., 1998). Yet, these usual visual motion areas did not respond selectively to the 3D-motion attribute. What could be the neural substrate for processing the 3D-motion attribute?

#### *Inferior parietal lobule (IPL)*

Even if the exact localisation of the foci slightly varied from stimulus-driven to attention-related conditions, we systematically found activity related to the 3D-motion attribute in the inferior parietal lobule. Moreover, no significant differences were found in this region between conditions (unattended transitions, task or attended transitions).

It has already been proposed that the IPL could mediate a high-level motion analysis based on saliency, independent of the usual luminance-based system of the SFM network (Claeys et al., 2003). The inferior parietal lobule is consistently involved in the perception of motion direction in paradigms using a variety of visual and audiovisual motion stimuli (Shulman et al., 1999; Claeys et al., 2003; Luks and Simpson, 2004; Baumann and Greenlee, 2007). The present results confirm the involvement of the IPL in processing the perceived (3D) motion of objects, which should be distinguished from the processing of the 2D-visual-speed distribution.

#### *Lateral temporal cortex: motion information storage and motion coherence*

A left middle temporal area was found when comparing selective attention to 3D motion with selective attention to 3D form. This region, anterior and superior to V5+, is not commonly found in the visual literature, particularly for the left hemisphere. When present, the activity of the middle temporal gyrus seems mostly associated with activity of the superior temporal sulcus (STS) and both are found in the perception of biological motion and human actions (Bonda et al., 1996; Decety and Grezes, 1999).

The middle temporal gyrus (MTG) is specifically activated by the perception of tool motion (Beauchamp et al., 2002). It could be involved in the analysis of movement intention as suggested by Grezes et al. (1999) or it could store information about the motion of manipulable objects (Beauchamp et al., 2002; Chao et al., 2002). In our case, the activity of the MTG might be explained by the storage of a reference 3D motion. Indeed, six participants reported that they memorized the current direction of motion in order to ensure they would not miss a transition. The MTG activity could also be related to the anticipation of the direction of motion expected by the observer.

Besides the MTG, the STS has been found when comparing coherent motion or texture patterns with incoherent ones (Braddick et al., 2000; Braddick et al., 2001). This is interesting if we consider that the perception of global 3D motion relies on the coherence of the speed distribution, whereas the perception of the structure relies on subtle variations of the speed distribution due to motion parallax. Thus, the fact that the STS was more activated when subjects were attending to 3D motion than when they were attending to form is consistent with the observers selecting spatially coherent information to perform the task.

*Overall, the usual visual motion areas (including V2, V3a, V5+ and the posterior IPS) do not seem to underlie the perception of the 3D motion attribute. They are more likely involved in processing the 2D motion input that can be used for both 3D motion and form perception. In contrast, the perception of 3D motion appears to rely on a high-level multimodal system of motion analysis encompassing IPL and STS regions.*

### ***Stimulus-driven vs. goal-directed***

In the present study, we used both stimulus-driven transitions and endogenous attention to modulate the activity related to the 3D form or 3D motion processing. After previous studies by Corbetta et al. (1991) or Huk and Heeger (2000), we were expecting that paying attention to 3D form or 3D motion would enhance the activity in cortical areas processing the corresponding unattended transitions. However, although both attended and unattended transitions revealed a segregation between the ventral and dorsal pathways, the precise areas

recruited by the attended transitions were functionally different from those recruited by the unattended transitions. How can we interpret this result?

*A matter of spatial attentional control?*

Corbetta and Shulman (2002) described two fronto-parietal pathways that can be recruited when the experimental situation requires the observers to shift their attention in space. The dorsal attentional pathway, encompassing the posterior parietal cortex (intraparietal sulcus and superior parietal lobule) and the frontal eye field, participates in the goal-directed control of attention. The ventral attentional pathway, encompassing the temporo-parietal junction (IPL and superior temporal gyrus) and the ventral frontal cortex (inferior and middle frontal gyrus), is involved when the orienting of spatial attention is driven by the stimulation. Although the present study was designed to avoid spatial shifts of attention, the movement of our stimulus may have oriented the participants' attention along the motion direction. Could stimulus-driven versus goal-directed shifts of attention account for the differences between unattended and attended transitions?

The posterior IPS, which could match the putative goal-directed network, is found in none of our main contrasts. The frontal focus that is found in the contrast for unattended transitions is far superior to what would be expected for the stimulus-driven attentional network. In contrast, the set of dorsal areas highlighted by the contrast for unattended transitions (parieto-frontal and superior temporal sulcus) may correspond better to areas described by Hopfinger et al. (2000) involved in the top-down control of spatial attention. Thus, neither the results for the attended transitions nor the results for the unattended transitions fit the spatial attention networks. We conclude that, although our experimental conditions may require a control of spatial attention, the present data reveal neither clear nor significant difference of attentional control between the form and 3D-motion attributes.

*A matter of attribute complexity?*

Recent data show that the modulation of the baseline activity related to attentional selection of a visual dimension, between stimulus presentations, can be independent of the response to stimulus onset (McMains *et al.*, 2007). Our stimuli, however, were presented without gap between the attribute transitions. In such conditions, Chawla et al. (1999) found that selective attention to stimulus motion or to stimulus colour induced congruent variations of the attention-related and stimulus-driven activity in areas V5 and V4. Why is not a similar congruence found for the 3D-motion and form attributes?

Up to now, the effect of the selective attention to features on the activity of feature-specific areas has been shown using simple dimensions, such as 2D motion, colour and shape, corresponding to independent visual cues (Corbetta et al., 1991; Chawla et al., 1999). It is possible that the modulatory effect observed in visual areas such as V4 and V5 be related to the visual cue from which the object attributes are extracted. In the present study, SFM perception allowed us to dissociate the processing of perceptual attributes from the processing of visual cues. Indeed, the 3D motion and form attributes we used are both extracted from the same visual cue, which is motion distribution. This shared origin could explain why a similar enhancement of V5 was observed when participants attended to either form or 3D motion. This may also explain why most participants were able to perceive form transitions during the motion task and motion transitions during the form task, while they hardly perceived colour transitions during these tasks.

In the ventral pathway, the attended transitions seem to induce activity in later, more specialized, shape sensitive areas than the unattended transitions. The observers' attention may therefore gate the level until which the complex attributes of the visual stimulus are processed and built-up. For independent dimensions such as colour and 2D visual motion, the perceptual build-up required to perform the task may be more limited, which could explain why attentional selection does not allow segregating different stages of the processing.

#### *Activation and deactivation*

In the present study, unattended transitions mainly induced enhanced activity. In contrast, attended transitions induced a large decrease of activity in the competing pathway. The suppression of unattended stimuli has already been described in human sensory cortices with paradigms modulating spatial attention (Slotnick et al., 2003; Muller et al., 2008) or cross modal attention (Ghatan et al., 1998; Johnson and Zatorre, 2005). Several authors also reported feature-related decreased activity in early visual areas. For instance, Pollmann et al. (2000) found decreased activity in the striate cortex when subjects shifted attention from one visual dimension to another. Sterzer and Kleinschmidt (2005) found decreased activity in V1 when subjects' percept did not follow the feature changes of the stimulus. These results have generally been interpreted as a suppression of the sensory entries for dimensions irrelevant to the perceptual decision.

The present data however suggest that feature-related suppression also occurs in higher-tier visual areas. A recent study by Nobre et al (2006) reports late electrophysiological effects of feature-related attention, presumably occurring in the fusiform gyrus, in the context of negative priming. The comparison with fMRI results however is not straightforward and there

is no evidence that the modulation of evoked potential by negative priming would result in a decreased BOLD activation. To our knowledge no fMRI data are available showing suppressive effects related to feature selection in high-level visual areas.

To summarize, our results are consistent with two coexisting mechanisms: presumably excitatory and inhibitory. In the context of an incidental task, visual transitions of the 3D stimulus engage the first stages of the form build-up in the SFM network, which would correspond to excitatory mechanisms. Inhibitory mechanisms may occur in the context of active detection to select the attribute of interest, thus decreasing the activity in areas processing the competing attribute.

## CONCLUSION

We distinguished the neural bases involved in processing 3D form and 3D motion attributes during structure from motion perception. It is the first time, to our knowledge, that the classical segregation between the ventral and dorsal pathways is highlighted so clearly. BOLD activity related to the form attribute was found in expected regions of the ventral pathway, including the lateral occipital and ventral temporal cortices. The processing of the 3D motion attribute however did not selectively involve the expected “visual motion areas” such as hMT/V5+, suggesting that the perception of the 3D motion direction is not exclusively processed within these areas. Instead, the 3D motion attribute specifically involved “high-level motion” areas located in the inferior parietal lobule and the superior temporal sulcus.

We were also able to segregate the regions subtending the visual processing of 3D motion and form, when these are not attended, from the substrate subtending their endogenous attentional selection. We found two functionally different substrates, contrary to what was previously showed with simple visual dimensions. We conclude that paradigms modulating the attention and paradigms modulating the stimulus do not necessarily provide similar evidence of function localization.

Finally, we propose that the attentional selection of simple visual dimensions and complex attributes operate at different stages of perceptual processing. While attention to form or attention to 3D motion may enhance the activity of the areas processing the visual motion dimension from which they are extracted, our results suggest that the selection of one perceptual attribute involves transient decreases of activity in high-level areas of the competing visual stream.

## **ACKNOWLEDGMENT**

We wish to thank Jean Lorenceau, Nathalie George and Catherine Tallon-Baudry for helpful discussions and comments on the manuscript. This work was supported by the ACC-SV program (n° 951261/12 from the French Ministère de la recherche) and IFR49. Scanning was conducted at the Service Hospitalier Frédéric Joliot, CEA, with the help of Pierre-Gilles Henri. ALP was supported in part by a fellowship from CEA, IFSBM, and a grant of neuro-ophthalmology from la Fondation de France.

## REFERENCES

- Baumann, O. and Greenlee, M. W. 2007. Neural Correlates of Coherent Audiovisual motion Perception. *Cereb Cortex*. **17**: 1433-1443.
- Beauchamp, M. S., Lee, K. E., Haxby, J. V. and Martin, A. 2002. Parallel visual motion processing streams for manipulable objects and human movements. *Neuron*. **34**: 149-159.
- Bonda, E., Petrides, M., Ostry, D. and Evans, A. 1996. Specific Involvement of Human Parietal Systems and the Amygdala in the Perception of Biological motion. *J Neurosci*. **16**: 3737-3744.
- Braddick, O. J., Brien, J. M., Wattam, B., Atkinson, J., Hartley, T. and Turner, R. 2001. Brain areas sensitive to coherent visual motion. *Perception*. **30**: 61-72.
- Braddick, O. J., Brien, J. M., Wattam, B., Atkinson, J. and Turner, R. 2000. form and motion coherence activate independent, but not dorsal/ventral segregated, networks in the human brain. *Current biology*. **10**: 731-734.
- Braunstein, M. L. and Andersen, G. J. 1984. Shape and depth perception from parallel projections of three-dimensional motion. *J Exp Psychol Hum Percept Perform*. **10**: 749-760.
- Brett M, Anton J-L, Valabregue R & Poline J-B (2002). Region of interest analysis using an SPM toolbox [abstract]. 8th International Conference on Functional Mapping of the Human Brain, June 2-6, 2002, Sendai, Japan. *Neuroimage*. **16**: 2.
- Büchel, C., Josephs, O., Rees, G., Turner, R., Frith, C. and Friston, K. 1998. The functional anatomy of attention to visual motion. A functional MRI study. *Brain*. **121**: 1281-1294.
- Chao, L. L., Weisberg, J. and Martin, A. 2002. Experience-dependent Modulation of Category-related Cortical Activity. *Cereb Cortex*. **12**: 545-551.
- Chawla, D., Rees, G. and Friston, K. 1999. The physiological basis of attentional modulation in extrastriate visual areas. *Nat Neurosci*. **2**: 671-676.
- Claeys, K. G., Lindsey, D. T., De Schutter, E. and Orban, G. A. 2003. A higher order motion region in human inferior parietal lobule: evidence from fMRI. *Neuron*. **40**: 631-642.
- Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P. and Shulman, G. L. 2000. Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nat Neurosci*. **3**: 292-297.

- Corbetta, M., Miezin, F., Shulman, G. and Petersen, S. 1991. Selective attention modulates extrastriate visual regions in humans during visual feature discrimination and recognition. *Ciba Found Symp.* **163**: 165-175.
- Corbetta, M. and Shulman, G. L. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci.* **3**: 201-215.
- Cornette, L., Dupont, P., Rosier, A., Sunaert, S., Van Hecke, P., Michiels, J., Mortelmans, L. and Orban, G. 1998. Human brain regions involved in direction discrimination. *J Neurophysiol.* **79**: 2749-2765.
- Cornilleau-Pérès, V. and Droulez, J. 1994. The Visual Perception of Three-Dimensional Shape from Self-motion and Object-motion. *Vision Research.* **34**: 2331-2336.
- Decety, J. and Grezes, J. 1999. Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Sciences.* **3**: 172-178.
- Denys, K., Vanduffel, W., Fize, D., Nelissen, K., Peuskens, H., Van Essen, D. and Orban, G. A. 2004. The processing of visual shape in the cerebral cortex of human and nonhuman primates: a functional magnetic resonance imaging study. *J Neurosci.* **24**: 2551-2565.
- Dupont, P., Orban, G. A., De Bruyn, B., Verbruggen, A. and Mortelmans, L. 1994. Many areas in the human brain respond to visual motion. *J Neurophysiol.* **72**: 1420-1424.
- Freire, L. and Mangin, J. F. 2001. motion correction algorithms may create spurious brain activations in the absence of subject motion. *Neuroimage.* **14**: 709-722.
- Ghatan, P. H., Hsieh, J. C., Petersson, K. M., Stone-Elander, S. and Ingvar, M. 1998. Coexistence of attention-based facilitation and inhibition in the human cortex. *Neuroimage.* **7**: 23-29.
- Goodale, M. 1998. Vision for perception and vision for action in the primate brain. *Proceedings of the Novartis Found Symp.* **218**: pp. 21-34.
- Greenlee, M. 2000. Human cortical areas underlying the perception of optic flow: brain imaging studies. *International review of neurobiology* **44**:269-292.
- Grezes, J., Costes, N. and Decety, J. 1999. The effects of learning and intention on the neural network involved in the perception of meaningless actions. *Brain.* **122**: 1875-1887.
- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y. and Malach, R. 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron.* **24**: 187-203.



- Grill-Spector, K., Kushnir, T., Edelman, S., Itzhak, Y. and Malach, R. 1998. Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron*. **21**: 191-202.
- Grosbras, M. H., Leonards, U., Lobel, E., Poline, J. B., LeBihan, D. and Berthoz, A. 2001. Human cortical networks for new and familiar sequences of saccades. *Cereb Cortex*. **11**: 936-945.
- Hopfinger, J. B., Buonocore, M. H. and Mangun, G. R. 2000. The neural mechanisms of top-down attentional control. *Nat Neurosci*. **3**, 284-291.
- Huk, A. and Heeger, D. 2000. Task-related modulation of visual cortex. *J Neurophysiol*. **83**: 3525-3536.
- Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L. and Haxby, J. V. 1999. Distributed representation of objects in the human ventral visual pathway. *Proc Natl Acad Sci USA*. **96**: 9379-9384.
- Jezzard, P. and Balaban, R. S. 1995. Correction for geometric distortion in echo planar images from B0 field variations. *Magn Reson Med*. **34**: 65-73.
- Johnson, J. A., Zatorre, R. J. 2005. Attention to simultaneous unrelated auditory and visual events: behavioral and neural correlates. *Cerebral cortex (New York, NY : 1991)* **15**:1609-1620.
- Kourtzi, Z., Erb, M., Grodd, W. and Bulthoff, H. H. 2003. Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cereb Cortex*. **13**: 911-920.
- Kriegeskorte, N., Sorger, B., Naumer, M., Schwarzbach, J., van den Boogert, E., Hussy, W. and Goebel, R. 2003. Human cortical object recognition from a visual motion flowfield. *J Neurosci*. **23**: 1451-1463.
- Longuet-Higgins, H. C. and Prazdny, K. 1980. The interpretation of a moving retinal image. *Proc R Soc Lond B Biol Sci*. **208**: 385-397.
- Luks, T. L. and Simpson, G. V. 2004. Preparatory deployment of attention to motion activates higher-order motion-processing brain regions. *Neuroimage*. **22**: 1515-1522.
- Martin, A., Wiggs, C. L., Ungerleider, L. G. and Haxby, J. V. 1996. Neural correlates of category-specific knowledge. *Nature*. **379**: 649-652.
- McMains, S. A., Fehd, H. M., Emmanouil, T. A., Kastner, S. 2007. Mechanisms of feature- and space-based attention: response modulation and baseline increases. *Journal of neurophysiology* **98**:2110-2121.

- Muller, N. G., Ebeling, D. 2008. Attention-modulated activity in visual cortex--more than a simple 'spotlight'. *NeuroImage* **40**:818-827.
- Murray, S. O., Olshausen, B. A. and Woods, D. L. 2003. Processing Shape, motion and Three-dimensional Shape-from-motion in the Human Cortex. *Cereb Cortex*. **13**: 508-516.
- Newell F. N., Wallraven C., Huber S. 2004. The role of characteristic motion in object categorization. *Journal of vision*. **4**:118-129.
- Nobre, A. C., Rao, A., Chelazzi, L. 2006. Selective attention to specific features within objects: behavioral and electrophysiological evidence. *Journal of cognitive neuroscience* **18**:539-561.
- Orban, G. A., Sinaert, S., Todd, J. T., Van Hecke, P. and Marchal, G. 1999. Human cortical regions involved in extracting depth from motion. *Neuron*. **24**: 929-940.
- Paradis, A.-L., Cornilleau-Pérès, V., Droulez, J., Van de Moortele, P.-F., Lobel, E., Berthoz, A., Bihan, D. L. and Poline, J.-B. 2000. Visual perception of motion and 3D structure-from-motion: an fMRI study. *Cereb Cortex*. **10**: 772-783.
- Paradis, A.-L., Droulez, J., Cornilleau-Pérès, V. and Poline, J.-B. 2001. Neural bases of the perception of 3-D structure-from-motion: insights from event-related fMRI and attention modulation [Abstract]. Proceedings of the ECVP, Kusadasi, Turkey. *Perception*, **30** supplement.
- Peuskens, H., Claeys, K. G., Todd, J. T., Norman, J. F., Van Hecke, P. and Orban, G. A. 2004. Attention to 3-D Shape, 3-D motion, and Texture in 3-D Structure-from-motion Displays. *J Cogn Neurosci*. **16**: 665-682.
- Pollmann, S., Weidner, R., Muller, H. J., von Cramon, D. Y. 2000. A fronto-posterior network involved in visual dimension changes. *Journal of cognitive neuroscience* **12**:480-494.
- Rogers, B. and Graham, M. 1979. motion parallax as an independent cue for depth perception. *Perception*. **8**, 125-134.
- Shulman, G. L., Ollinger, J. M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Petersen, S. E. and Corbetta, M. 1999. Areas involved in encoding and applying directional expectations to moving objects. *J Neurosci*. **19**: 9480-9496.
- Slotnick, S. D., Schwarzbach, J. and Yantis, S. 2003. Attentional inhibition of visual processing in human striate and extrastriate cortex. *Neuroimage*. **19**: 1602-1611.

- Sterzer, P., Kleinschmidt, A. 2005. A neural signature of colour and luminance correspondence in bistable apparent motion. *The European journal of neuroscience* **21**:3097-3106.
- Stone, J. V. 1998. Object recognition using spatiotemporal signatures. *Vision Res* **38**(7): 947-51.
- Sunaert, S., Van Hecke, P., Marchal, G. and Orban, G. A. 2000. Attention to speed of motion, speed discrimination, and task difficulty: an fMRI study. *Neuroimage*. **11**: 612-623.
- Tootell, R. B. H., Mendola, J. D., Hadjikhani, N. K., Ledden, P. J., Liu, A. K., Reppas, J. B., Sereno, M. I. and Dale, A. M. 1997. Functional analysis of V3A and related areas in human visual cortex. *J Neurosci*. **17**: 7060-7078.
- Treue, S. and Maunsell, J. H. R. 1996. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*. **382**: 539-541.
- Ungerleider, L. and Mishkin, M. 1982. Two cortical visual pathways. In *Analysis of visual behaviour*, Ingle, D. J. et al. (Eds.). MIT Press, Cambridge, MA, pp. 549-586.
- Valyear, K. F., Culham, J. C., Sharif, N., Westwood, D., Goodale, M. A. 2006. A double dissociation between sensitivity to changes in object identity and object orientation in the ventral and dorsal visual streams: a human fMRI study. *Neuropsychologia* **44**:218-228.
- Xiao, D. K., Marcar, V. L., Raiguel, S. E. and Orban, G. A. 1997. Selectivity of Macaque MT/V5 Neurons for Surface Orientation in Depth Specified by motion. *Eur J Neurosci*. **9**, 956-964.

## FIGURES



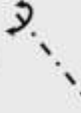


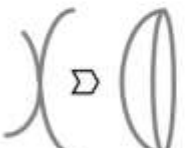

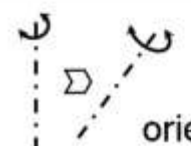

Task Factors Transitions	Detect changes of ...		
	...colour 	...form 	...motion 
 colour	- 	Same transitions: Contrasts of attention-related activity (2)	
 form	Same task, unattended transitions: Contrasts of stimulus-driven activity (1)		- 
 motion orientation	Attended transitions: Contrasts of stimulus-driven & attention-related activity (3)		- 

Figure 1: Experimental conditions and their relation to the contrasts of interest. Overall six contrasts were calculated to compare the neural substrate of the 3D form and motion processing (2 opposite contrasts) when the activity was stimulus-driven (1); attention-driven (2) or driven jointly by the stimulus and the endogenous attention (3).

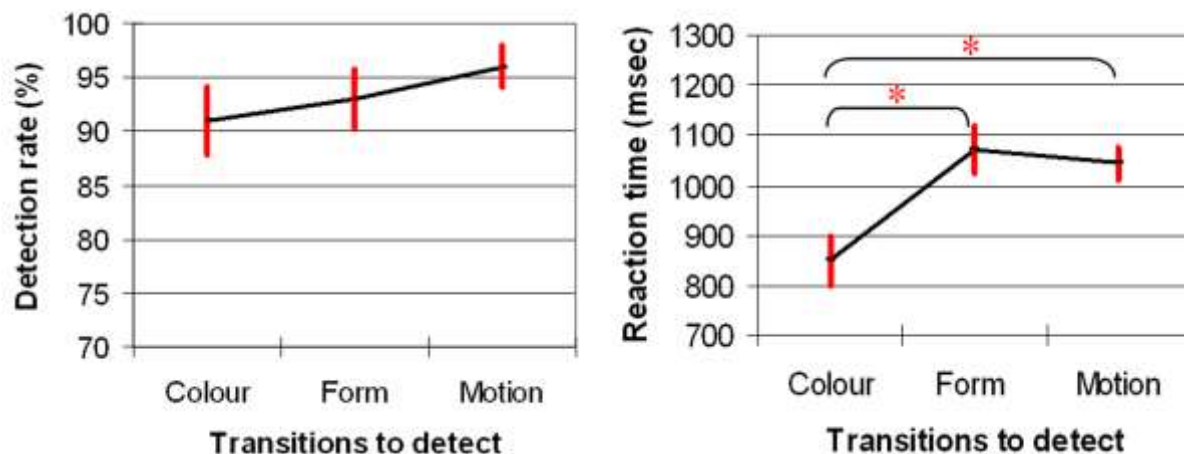


Figure 2: Mean detection rate and reaction times for the 3 types of transitions to be detected ( $\pm$  standard error, 10 subjects, \* indicate significant differences).

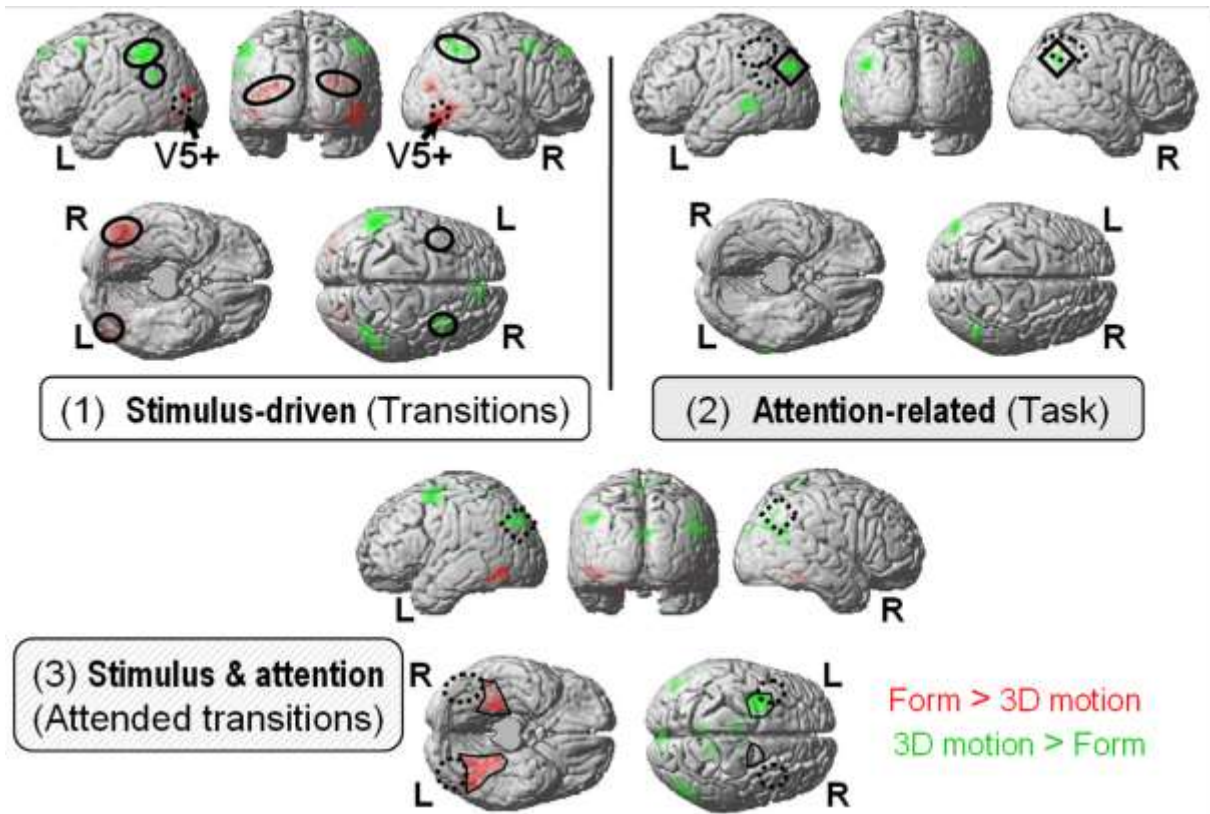


Figure 3: Contrasts between form and 3D motion processing for three conditions of activation. In red: form-related activity > 3D motion related activity; in green: 3D motion related activity > form-related activity. Solid figures highlight locations of interest in the current contrast. Dotted figures remind locations found in other conditions. (1) Contrasts of unattended transitions. Dotted circles indicate the position of V5+, as localized from an independent experiment with the same subjects. The solid circle on the left superior frontal sulcus encompasses a focus of maximal local activity, which did not pass the cluster threshold. (2) Contrasts of tasks. No activity is found for Form minus 3D motion. In the inferior parietal lobule (IPL), the areas preferentially activated by the attention to the 3D motion (solid diamond) are slightly posterior to that elicited by the unattended 3D-motion transitions (dotted circles). (3) Contrasts of attended transitions. The present IPL focus overlaps the attention-related one (dotted diamonds), but the frontal and infero-temporal foci differ from those found in the contrasts of unattended transitions (dotted circles). L = left hemisphere; R = right hemisphere.

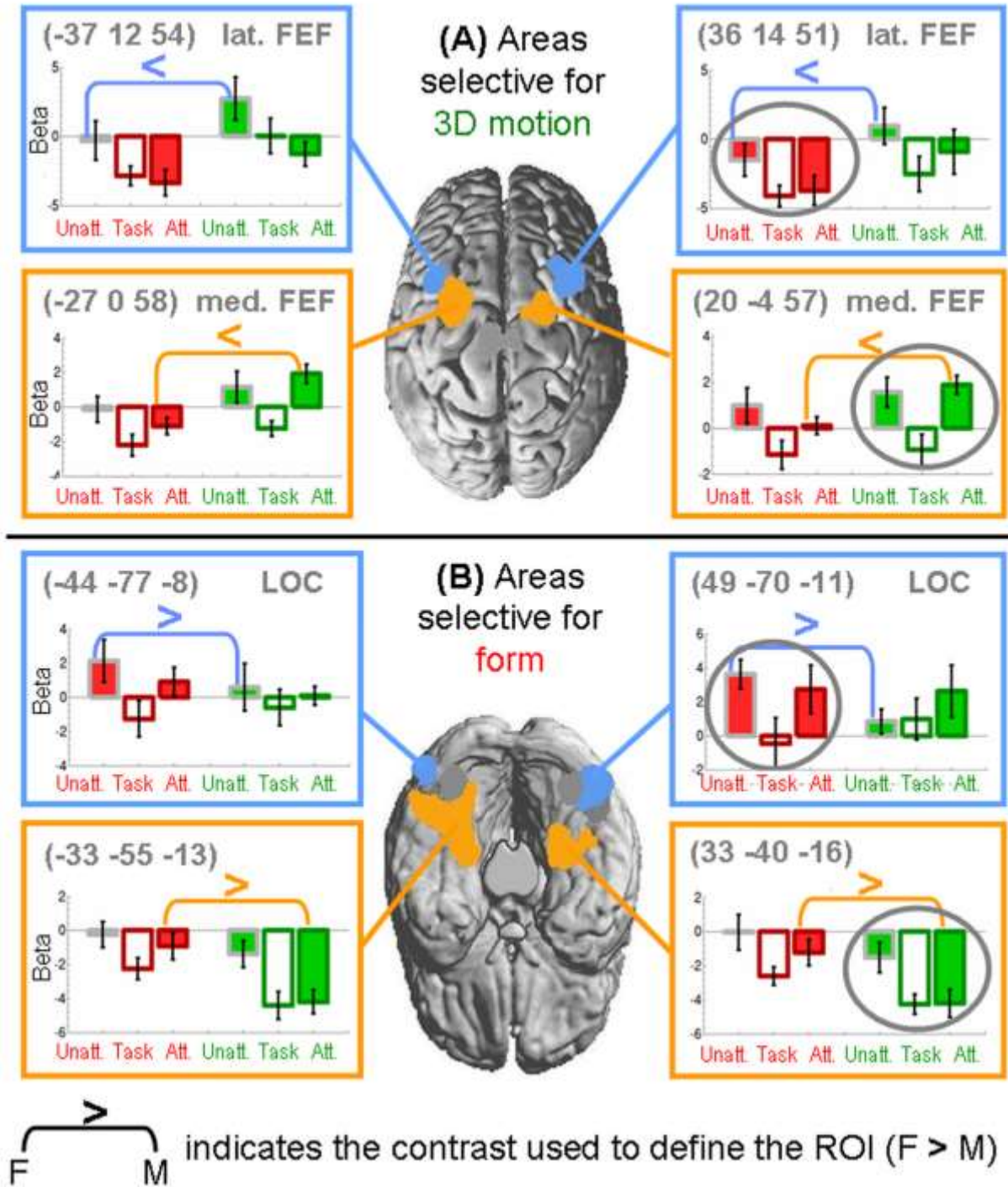


Figure 4: Detailed activity of selected areas relative to a low level baseline. Each bar represents the response of the region (Beta) to a stimulus transition (averaged over 10 subjects). Activity related to the form attribute is plotted in red; activity related to the 3D motion attribute is plotted in green. From left to right, we find the responses to the unattended transitions of the attribute (Unatt.); the response to the colour transitions occurring while subjects attended to the attribute (Task); and the response to the attended transitions of the attribute (Att.).



(A) shows the activity pattern of dorsal areas selective for 3D motion (regions of interest represented on the superior aspect of the brain). (B) shows the activity pattern of ventral areas selective for the form attribute (regions of interest represented on the inferior aspect of the brain). The coordinates in parentheses correspond to the centres of mass of the regions. Grey circles underline the attribute inducing the most significant responses (either positive or negative) relative to the low level baseline. From these responses, we could determined that the functional patterns of activity differed between the areas delineated by a contrast of attended transition (in orange) and the areas delineated by a contrast of unattended transitions (in blue) (see text).

### Supplementary Material

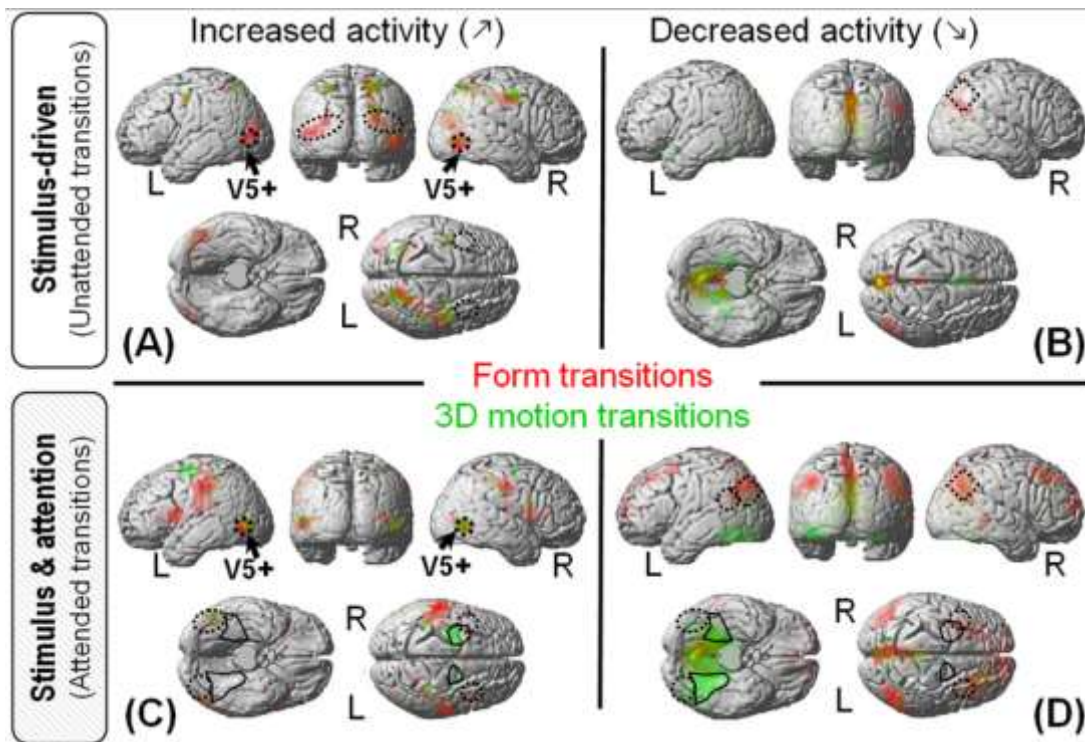


Figure 5: Maps of increased and decreased BOLD activity relative to a low level baseline, for the transitions of form and 3D motion. In red, the activity related to the form transitions; in green, the activity related to the 3D motion transitions. (A) and (B) respectively show the regions of increased and decreased activity for the unattended transitions. (C) and (D) respectively show the regions of increased and decreased activity for the attended transitions. Dotted circles delineate the regions of interest from the contrasts of unattended transitions; dotted diamonds delineates the regions of interest from the contrasts of tasks (see Figure 3). Solid line drawings delineate ventral and dorsal areas that were highlighted by the contrasts of attended transitions.

## TABLES

**Table 1: Activity foci in the regions elicited by the contrasts of stimulus-driven activity**

Region		z- score	Coordinates (mm)		
			x	y	z
<u>A/ preferential response to the unattended <i>form</i> transitions</u> <u>(p&lt;0.001)</u>					
<i>Superior occipital s.</i> (V3A)	<i>Left</i>	<i>3.67</i>	<i>-15</i>	<i>-93</i>	<i>12</i>
	<i>Right</i>	<i>3.31</i>	<i>24</i>	<i>-81</i>	<i>15</i>
	<i>Right</i>	<i>3.03</i>	<i>24</i>	<i>-75</i>	<i>24</i>
<i>Lateral occipital s.</i>	<i>Left</i>	<i>3.47</i>	<i>-36</i>	<i>-84</i>	<i>9</i>
	<i>Left</i>	<i>3.08</i>	<i>-27</i>	<i>-90</i>	<i>15</i>
	<i>Right</i>	<i>3.30</i>	<i>39</i>	<i>-72</i>	<i>12</i>
<i>Inferior temporal s.</i> (LOC)	<i>Left</i>	<i>3.13</i>	<i>-39</i>	<i>-63</i>	<i>-6</i>
	<i>Right</i>	<i>4.06</i>	<i>51</i>	<i>-57</i>	<i>-3</i>
	<i>Right</i>	<i>3.02</i>	<i>54</i>	<i>-72</i>	<i>-12</i>
<i>Temporo-occipital s.</i>	<i>Right</i>	<i>3.00</i>	<i>33</i>	<i>-66</i>	<i>-18</i>
<i>Collateral s.*</i> (p <sub>cluster</sub> = 0.058)	<i>Right</i>	<i>3.09</i>	<i>18</i>	<i>-69</i>	<i>-21</i>
<u>B/ preferential response to the unattended <i>3D-motion</i> transitions (p&lt;0.001)</u>					
Inferior parietal lobule (supramarginal g.)	Left	3.61	-51	-48	48
	Left	3.06	-57	-39	45
	Right	3.17	45	-57	48
Lateral TPJ	Left	3.25	-60	-54	27
Superior frontal g. (BA8/9)	Mesial	3.29	6	42	48
Pre-central s.	Right	3.02	33	21	51
Middle frontal g.	Right	3.52	36	9	57
(p <sub>cluster</sub> = 0.083)	Left*	3.73	-36	12	51

s. = sulcus; g. = gyrus; BA = brodmann area; LOC = Lateral occipital complex; Lateral TPJ = Lateral temporo-parietal junction (posterior end of the sylvian fissure); \* in grey, areas under the cluster threshold  $p = 0.05$ ; *in italic*, foci belonging to the SFM network (i.e. activated by SFM stimuli).



**Table 2: Activity foci in the regions elicited by the attention-related contrasts**

Region		z-score	Coordinates (mm)		
			<i>x</i>	<i>y</i>	<i>z</i>
<u>Preferential response during the <i>3D-motion</i> task (p&lt;0.001)</u>					
MTG	Left	3.72	-72	-42	-3
Inferior parietal lobule	Left	3.61	-45	-78	33
(angular gyrus)	Right	3.56	51	-57	42
MTG = middle temporal gyrus					

**Table 3: Activity foci in the regions elicited by the contrast combining attention and stimulus effects**

Region		z- score	Coordinates (mm)		
			<i>x</i>	<i>y</i>	<i>z</i>
<u>A/ preferential response to the attended <i>form</i> transitions</u>					
<u>(p&lt;0.001)</u>					
<i>Inferior temporal s.</i>	<i>Left</i>	2.97	-36	-69	-9
<i>Temporo-occipital s.</i>	<i>Left</i>	2.99	-45	-60	-9
	<i>Right</i>	3.61	45	-42	-15
Collateral s.	Left	3.14	-27	-45	-15
	Left	3.01	-21	-36	-21
	<i>Right</i>	3.34	33	-36	-18
Insula	Left	3.65	-30	6	12
<u>B/ preferential response to the attended <i>3D-motion</i></u>					
<u>transitions (p&lt;0.001)</u>					
Cuneus		3.74	3	-90	24
Superior temporal s.	Right	3.01	48	-54	12
Inferior parietal lobule (angular g.)	Left	3.07	-42	-81	39
	Right	3.04	48	-75	30
Pre-central s. / Superior frontal s.	Left	3.23	-27	-3	57
	Left	3.33	-18	-6	54
	Left	3.07	-24	6	57
	(p <sub>cluster</sub> = 0.056) Right*	3.85	21	-3	57
Posterior cingulate s. (marginal segment)	Left	3.15	-3	-51	66
	Right	3.52	12	-45	66
Paracentral lobule	Left	3.73	-12	-21	60
Caudate nucleus	Left	3.40	-18	12	18

g. = gyrus; s. = sulcus; in the same cluster, only foci more than 9 mm apart are reported; \* in grey, areas under the cluster threshold  $p=0.05$ ; in *italic*, foci belonging to the SFM network (i.e. activated by SFM stimuli)

## SUPPLEMENTARY MATERIAL

To better understand why the contrasts between attended transitions differed from the contrasts between unattended transitions, we calculated the maps of increased or decreased activity with respect to a low level baseline in response to these transitions. We describe here the global activation of these maps in relation with the SFM network. We also detail the activity of two areas that were activated in contrasts of unattended transitions but had no obvious counterpart in the contrasts of attended transitions.

### *Maps of enhanced and decreased activity at the attended and unattended transitions*

#### *Unattended transitions*

The patterns of increased activity elicited by the unattended transitions of form and 3D motion both tightly match the SFM network (see Figure 5-A). Our results thus replicate in the controlled context of the incidental colour task, results previously obtained in passive viewing or while the attention was directed toward the 3D object.

Within this network, we have found that the superior occipital cortex and the infero-temporal sulcus (dominantly red in Figure 5-A) are selective for the form attribute (see Figure 3-1). By contrast, the intraparietal sulcus and the lateral occipital sulcus (hMT/V5+) are similarly activated by the transitions of both attributes (see the common activation in Figure 5-A and the absence of significant contrast in Figure 3-1).

The unattended transitions also induce decreased activity (see Figure 5-B). This decreased activity is mainly found in mesial structures (thalamus, hippocampus, cuneus, precuneus and cingulate) where it is common to the form and 3D-motion transitions. Only the inferior parietal lobule (angular gyrus) reveals a decrease of activity specific to the transitions of form.

#### *Attended transitions*

Contrary to the unattended transitions, the attended transitions only activate a small part of the SFM network: hMT/V5+ and the anterior end of the post-central sulcus, excluding the whole intraparietal sulcus (see Figure 5-C). According to the contrasts between attributes, this part of the SFM network is neither form selective nor 3D-motion selective (see Figure 3-3). Foci of enhanced activity also appear outside the SFM network. These are common to the form and 3D-motion transitions, as well. They correspond to the SMA and the primary motor cortex, which are probably related to the manual response activity; and the inferior/superior parietal lobule, which is possibly related to target detection (Corbetta *et al.*, 2000). Overall, the regions of increased activity do not appear in the contrasts of attended transitions.

In contrast, we find decreased activity bilaterally in the inferior parietal lobule (see dotted circles in Figure 5-D). This cluster overlaps the region of decreased activity at the unattended transitions (Figure 5-B), and closely corresponds to the 3D-motion selective regions found in the contrast of attended transitions. The inferior parietal lobule thus reveals decreased activity in response to all form transitions. Decreased activity is also found at the attended transitions of form along the superior frontal sulcus. This corresponds to the 3D-motion selective areas found in the contrast of attended transitions. Finally, extended decreased activity is found in the inferior temporal cortex, which largely overlaps the form-selective areas in the contrast of attended transitions.

*When attended, the transitions of form and 3D motion activate only a small part of the SFM network, but induce a spatially extended decrease of activity. Overall, the contrast maps for the attended transitions are not well explained by the patterns of increased activity, but rather by the pattern of decreased activity.*

### ***Focus on a form-selective and a 3D-motion selective area***

In the main result section, we checked that the areas that were closely located in the contrast of attended transitions and the contrast of unattended transitions were functionally different. The supplementary results presented here assess the behaviour of two regions that were activated in a contrast of unattended transitions but had no obvious counterpart in the contrasts of attended transitions. We measured the response of these regions to the unattended transitions with respect to our low level baseline, in order to evaluate whether these responses were increased or decreased. We also assessed the activity of these regions during the task and the attended transitions, for both the form and the 3D-motion attribute. We particularly tested the difference of activity between task and attended transitions, in order to assess the effect of the attribute transition when the transition was relevant to the task (we report uncorrected p values from the MarsBar statistics).

#### *Superior occipital cortex*

The superior occipital areas show increased activity relative to the low level baseline at the unattended transitions of form ( $p=0.004$  on the left and  $p=0.001$  on the right). In contrast, they reveal no significant activity related to the tasks (either form or 3D motion).

The right cluster also shows increased activity at the attended transitions of form ( $p=0.035$  on the right); and the left area is significantly more activated by the attended transitions of form than by the form task at colour transitions ( $p=0.02$ ). Overall the activity of the superior occipital cortex seems to be driven primarily by the stimulus processing whatever the task

context is. The absence of this area in the contrast of attended transitions may have been due only to a threshold effect. It reveals however that the selectivity to the attribute is better highlighted in the context of an incidental task, which likely corresponds to an automatic processing of the stimulus.

*Superior frontal gyrus (mesial BA 8/9)*

The mesial activity focus of the superior frontal gyrus was defined by a significant difference between the unattended transitions of 3D motion and the unattended transitions of form. We observe that this is due to an increased activity at the unattended transitions of 3D motion. In contrast, we find no significant activity, either decreased or increased, at the attended transitions of 3D motion or during the motion task. The region however shows decreased activity during the form task ( $p=0.01$ ) and the attended transitions of form ( $p=0.008$ ). Furthermore, we find no significant differences of activity between the task and attended transitions, either for the form or the 3D motion attribute. Hence, this region, which was selective for the transitions of 3D motion during an incidental task, does not appear selective for any transitions when these are relevant to the task.

*The behaviour of the dorsal area presented here confirms that the selective attention to one attribute does not always enhance the activity in the areas processing this attribute. On the contrary, we observe that focusing attention toward 3D motion may annihilate the selectivity of the region to the visual transitions of this attribute.*